MEMORIAL UNIVERSITY OF NEWFOUNDLAND

DEPARTMENT OF MATHEMATICS AND STATISTICS

A2 **MATH 6205** DUE: FEB. 2ND, 2022

Upload your solution to Brightspace at the beginning of class on Wednesday Feb. 2nd, 2022.

1. We have seen in class that selecting a machine learning algorithm is only one step in completing a machine learning project. Besides building a data pipeline, a central further component is hyperparameter tuning for the selected learning algorithm. Proper hyperparameter tuning can boost the performance of an appropriate algorithm quite substantially and thus should be carried out before shipping the algorithm to a client.

   In this assignment, you will carry out hyperparameter tuning for the MNIST handwritten digit dataset, using both *grid search* and *random search* (see `sklearn` documentation).

   Complete the following tasks:

   (a) The baseline model for your task should be a SVM with a linear kernel. Train this model on your training data and evaluate the accuracy on the testing data.

   (b) To improve upon this baseline, try a SVM with a radial basis function (RBF) kernel. Since you will find that the RBF-SVM model performs better than the linear-SVM model, the former will be the proper candidate for our hyperparameter study.

   (c) Carry out a hyperparameter tuning study for the SVM model with RBF kernel. There are two parameters influencing the model, γ (the RBF shape parameter) and C (the regularization parameter). Both γ and C have default values for the given problem (read the `sklearn.svm.SVC` documentation to find/compute those default values).

   (d) For both grid search and random search set up a meaningful search space for your parameters, e.g. logarithmically spaced intervals containing the default parameters. Note that if your grid search or random search yields optimal hyperparameter values that are at the boundary of your search domains, then your search space may have been too narrow, and you will have to extend those search domains to make sure that you indeed found the optimal values for the parameters.

   (e) Report on the computational efficiency of grid search vs. random search, e.g. by clocking the total time it took to complete each algorithm to find suitable hyperparameter values.

   (f) Once you have found these optimal hyperparameters, train your models with them. Do the optimal parameters found by grid search or those found by random search give the best overall model?

   Note: Depending on your available computing hardware, grid search and random search on the full MNIST dataset may take up to one day or longer, which is not feasible for an assignment! Thus, it is perfectly acceptable to use only a subset of the training/test data available (e.g. 10% of the data should work reasonably well). Still, this complication illustrates a problem you'll likely face in practice when working on a real machine learning project: Your computational and time resources are not infinite and have to be an important consideration in selecting a proper algorithm for your problem (for the MNIST dataset other algorithms like convolutional neural networks are more appropriate).